

Tézisfüzet

Variációs módszerek a gépi látásban

MOLNÁR JÓZSEF

Témavezető: Prof. Csetverikov Dmitrij

Eötvös Loránd Tudományegyetem

Informatika Doktori Iskola

Az informatika alapjai és módszertana

A doktori program vezetője: Prof. Demetrovics János

Budapest 2011

I. Bevezetés

Az alábbi rövid bevezető összefoglalja a variációszámítás helyét tudományokban általában és speciálisan a gépi látásban. Az alapvető fogalmakat, amelyek későbbiekben gyakran előfordulnak dőlt betűkkel szedtem.

Egy diszciplína – amely a benne definiált objektumok nem triviális kapcsolatrendszerét (kölcsonhatásait) természetes módon matematikai fogalmakkal modellezi – általában alapegyenleteiben szintetizálódik. Ezek gyakran közönséges vagy parciális differenciálegyenlet-rendszerek, amelyek a diszciplína axiómáiból közvetlenül, vagy közvetett módon származnak. A közvetett módon, variációszámítási elvekből származtatott alapegyenleteket tartják matematikai szempontból a leginkább megalapozottnak. A variációs elvekből származtatott alapegyenletek ugyanis többet fejeznek ki a lokális kapcsolatoknál: a kölcsönhatások lokális jellegének megtartása mellett az egész rendszerre vonatkozó bizonyos elvek érvényesülését is garantálják. Ezek az elvek többnyire megmaradási, minimalizálási elvek. A módszer pedig amivel mindez biztosítható, *funkcionálok* szélsőérték helyének keresése. Ilyenek az elméleti fizika legkisebb hatás elvei mozgásegyenletek és téregyenletek levezetésére, vagy a görbült terek geometriájában az egyenes fogalmának általánosítása, a geodetikus, mint a tér két pontja közötti legrövidebb út.

A gépi látásban általában az *energia-minimalizálás analógia* vezet variációs elvek használatához. Ilyen pl. a képtartalom szegmentációja aktív kontúrral, ahol a szegmentáló görbe „energiáját” minimalizáljuk. Az energia a „külső” – képtartalom függő, és a „belső” – a görbe alakjától függő részenergiák összege. Hasonlóan az optikai áramlásban, ahol képsorozatok szomszédos képei közötti elmozdulásmező meghatározása a cél, egy „energiaminimalizáló” mezőt keresünk. Ez a mező kielégíti az optikai kényszer (pl. intenzitásállandóság) megmaradásának követelményét, a mező valamilyen belső jellemzőjének (pl. az elmozdulás vektorok divergenciája) minimalizálása mellett. Valamilyen mennyiség megmaradása ekvivalens változásának minimalizálásával ezért az „összes energia minimalizálása” értelmezés az utóbbi példára is helytálló. A gépi látásban felmerülő problémáknál gyakori tehát az energiaminimalizálási analógia használata¹. Jellemző továbbá a fenti példákkal illusztrált kettősség. Az „összenergiában” megjelenik a külső hatásokért felelős mennyiség: az *adattag*, és mellette valamilyen belső tulajdonság megőrzéséért felelős mennyiség: a *simasági tag*.

¹ Innen ered az irodalmában gyakran használt energia-funkcionál kifejezés.

A variációs számítás módszere a szélsőértékek meghatározására a funkcionálhoz rendelt *Euler-Lagrange egyenletek* származtatása. Az egyenletek típusa a probléma dimenziójától, az ismeretlen függvények számától és az ismeretlen függvények deriváltjainak rendjétől függ: a közönséges másodrendű differenciálegyenlettől a magasabb fokú parciális differenciálegyenlet rendszerekig tart. A parciális differenciálegyenletek numerikus megoldására iteratív módszerek ismertek, a megoldás az iteráció ú.n. fix pontja, amikor az egymást követő közelítések között a különbség küszöbérték alá csökken. Magasabb dimenziós térbe ágyazott sokaságokra végzett iterációs műveletet evolúciónak nevezzük. Az evolúciós módszerek közül sikerességével kiemelkedik a *Level Set* módszer.

Az alábbiakban összefoglaljuk a variációs módszerek alkalmazásának néhány tipikus területét, feladatukat a gépi látásban és a használt variációs módszereket. Az aktív kontúr, aktív felület (active contour, active surface) módszereket széles körben használják közvetlen képi információk szegmentálására (segmentation), de 3D objektumok és szintér rekonstrukciójára (reconstruction) is. A teljes variáció (total variation) módszer a képkorrekciós technikák (restoration) variációs alapú eljárása elmosódott, zajos képek javítására. A teljes variációs és aktív kontúros módszerek kombinációja egy lehetőség a képtartalom rekonstrukció (interpolation) variációs kezelésére, ahol a hiányzó képi információk pótlása történik a hiány környezetének adatai alapján úgy, hogy a fontos képi jellemzők, mint az élek és textúrák megjelenjenek az interpolált részleteken is. A variációs optikai áramlás (optical flow) alapvető módszer a képsorozatok szomszédos képei közötti mozgások elemzésének, de előfordul 3D szintér áramlás (scene flow) számítás is. Felhasználása sokrétű: videó tömörítési technológiák kulcsképek (key frame) alapján, robotika, gépjárművek asszisztens rendszerei, ember-gép interakciók. Optikai áramlási alapokon nyugszik a képregisztrációs problémák egy része is, ahol az alapp probléma a különböző szenzorok általi objektumrepresentációk illesztése (multispectral, multimodal registration) légi felvételeknél és az orvosi diagnosztikában.

II. A disszertáció felépítése

Egy rövid *Bevezetés* után a második fejezetben – *Variációs elvek, megjelenésük a gépi látásban* – szakirodalmi hivatkozásokkal számba vesszük a gépi látás azon területeit, ahol a variációs módszerek használata elterjedt. Néhány reprezentatív példán keresztül elemezzük funkcionálok jelentését, szerkezetüket, az adat és simasági tagok használatának módját. Ezeket a példákat a disszertáció későbbi fejezeteiben referenciaként is felhasználjuk. A Level Set formalizmus ismertetése után a variációszámítás módszereit mutatjuk be – alapeseteken keresztül – differenciálegyenletek származtatására. A fejezet zárásaként egy konkrét példán keresztül illusztráljuk az Euler-Lagrange egyenletek levezetését. A példa referenciaként szolgál az értekezés harmadik fejezetében.

A harmadik fejezet – *Optikai áramlás* – bevezetőjében bemutatjuk a módszer alkalmazási területeit, a módszerek egyfajta csoportosításával a variációs optikai áramlás jellemzőit, a kutatás motivációját (megvilágítás-változást tűrő alkalmazás) és a kapcsolódó kutatásokat. A fejezet második részében részletesen ismertetjük a *Keresztkorrelációs optikai áramlást*: a (nem centrális) normalizált keresztkorrelációs adattagot szürke árnyaltos és színes képekre, a közelítő Euler-Lagrange egyenleteket, a linearizálás és diszkretizálás elveit. A numerikus végeredményt összehasonlítjuk a bevezetésben példaként levezetett Horn-Schunck formulával. Az alfejezet szerves része az Euler-Lagrange egyenletek származtatása az **A melléklet**ben. A következő alfejezet a *Keresztkorrelációs optikai áramlás tesztsjelei* leírja a tesztkörülményekeket és a teszteredményeket csoportosítva szintetikus szürkeárnyaltos, kültéri és szintetikus színes képszekvenciákra. A fejezet összefoglalójában a módszer pontosságát hasonlítjuk a korszerű módszerekhez és a továbbfejlesztés lehetőségeit tárgyaljuk.

A negyedik fejezet – *Aktív kontúr* – bevezetőjében bemutatjuk az aktív kontúr alapú szegmentációs technikák fejlődését, típusait. A *Lokális régió alapú szegmentáció* alfejezetben bemutatjuk a motivációt (Optical Coherence Tomography technológiával készített felvételek rétegszegmentációja) és a bevezetett módszer elvárt tulajdonságait, amelyek alkalmassá tehetik a probléma kezelésére. A következő részben *Az alapmodell* kerül sor a legegyszerűbb lokális régió alapú modell részletes ismertetésére. Ez tartalmazza a szegmentáló görbe menti lokális régiók definícióját, a hozzárendelt energiafunkcionált, a származtatott normálirányú Euler-Lagrange egyenleteket (levezetésük a **B melléklet**ben található), a (közelítő) normáláramlási egyenletekhez rendelt Level Set egyenleteket és egy egyszerű statisztikai szeparátor függvényt. Az alfejezet az alpmódszer kritikájával zárul. *A modell finomításai*

részben a kétirányú továbbfejlesztés lehetőségeit és következményeit tárgyaljuk: a *másodfokú görbeközelítéssel* a lokális régiók mérete érintőirányban növelhető (robustusabb statisztika), míg az *optimális alakú integrálási tartomány* használatával a normálvektor irányú integrálási határok optimális megválasztásával a módszer hatékonysága fokozható (abban az értelemben, hogy kisebb átlagos különbségekre is használható). Az utóbbi továbbfejlesztésről bebizonyítjuk, hogy önmagában egy (lokális) variációsszámítási probléma. *A modell alkalmazása, eredmények* alfejezetben a tesztek körülményeit és eredményeit, egy lehetséges kétlépcsős technika alkalmazhatóságát tárgyaljuk a továbbfejlesztési lehetőségek (3D) számbavételével.

Az ötödik fejezet – *3D rekonstrukció* – bevezetőjében röviden összefoglaljuk a funkcionál-minimalizáláson alapuló 3D rekonstrukciós módszereket továbbá a leggyakrabban használt kameramodellt (lyukkamera modell); az ezen a modellen alapuló projektív és affín homográfiát (különböző nézetekből készített képek részletei közötti megfeleltetést), majd számba vesszük ezen megfeleltetések korlátait. Célul tűzzük ki másodrendű, a Level Set módszerrel kompatibilis másodrendű megfeleltetés levezetését. *A Lineáris transzformáció* részben a Level Set módszerrel kompatibilis lineáris transzformáció levezetését részletezzük, ennek lépései szolgálnak mintául a kvadratikus transzformáció levezetéséhez. *A Kvadratikus transzformáció* alfejezetben vezetjük le a képrészletek közötti kvadratikus transzformáció egyenleteit, amelyek mind a kamerák vetítési függvényeit, mind a megfigyelt felületet másodrendű differenciális mennyiségekkel közelítik. A levezetések részletei a C és D *mellékletekben* találhatók. Az alfejezetet a levezetett mennyiségek fix térbeli rácson való számításával zárjuk. Alternatív számítási módszert is megadunk az E *mellékletben*. *A kvadratikus transzformáció eredménye* alfejezetben elemezzük az eredményt, értelmezzük az egyes tagok jelentését, és összehasonlítjuk a lineáris (affín homográfia lyukkamera esetén) és kvadratikus megfeleltetések pontosságát/hatókörét. *A Kvadratikus transzformáció egy alkalmazása* záró alfejezetben ismertetjük a Feugas-Keriven által bevezetett többkamerás rekonstrukciós eljárást, amelyet a kvadratikus transzformáció validálására használtunk, továbbá a tesztkörülményeket és teszteredményeket.

A hatodik fejezet – *Tézisek* – a disszertáció téziseit sorolja föl. Az értekezésben használt jelölésrendszer a bevezetés előtt a *Jelölések* címsor, a hivatkozások az értekezés végén *Bibliográfia* címsor alatt találhatók.

III. A disszertációban tárgyalt új tudományos eredmények

A gépi látás három független területén végeztem kutatásokat. Ezek az *optikai áramlás*, az *aktív kontúr* és a *3D rekonstrukció*. Minden esetben nagy hangsúlyt fektettem a matematikai alapok tisztázására.



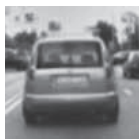
1. kép (mesterséges árnyékolással)



2. kép



Horn-Schunck



Keresztkorreláció



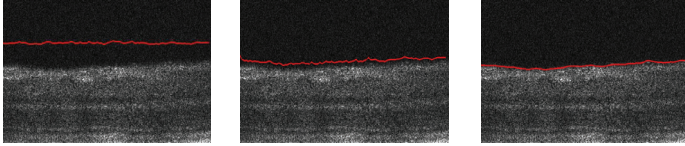
Horn-Schunck



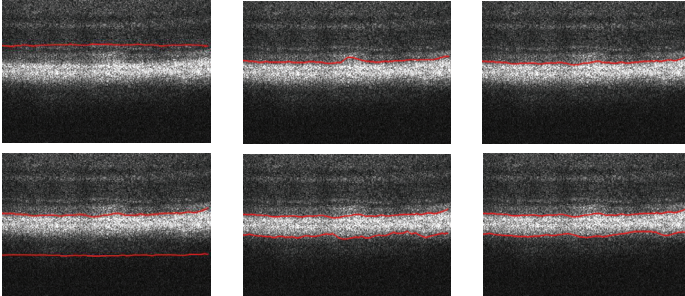
Keresztkorreláció

1. ábra : Kültéri videó két kockája (felül). Az optikai áramlással számított elmozdulásmezőt az első kép pixeleire alkalmazva a második kép rekonstrukcióját kapjuk. A kétféle módszerrel rekonstruált képrészletek az árnyékolt részletről és egy komplexebb területről (alul).

Az *optikai áramlás* esetében cél volt egy gyors, a színtér megvilágításbeli viszonyainak változását jól tűrő módszer kifejlesztése, amely alkalmas kültéri felvételek feldolgozására (1.ábra), és akár a színes megvilágításban beállt változások kezelésére is. Új eredmény született a normalizált keresztkorreláció adattagként való felhasználásából variációs keretekben: a Lagrange függvény speciális struktúrája (lokális integrálokból komponált mennyiség) következtében az Euler-Lagrange egyenleteket végtelen sorként származtattam. Többlépcsős linearizálással jól használható numerikus formulát adtam meg. Kifejlesztettem a numerikus formulát megvalósító szoftverkomponenst. Nagy gondot fordítottam a módszer tesztjeire, a teszteteket elvégeztem szintetikus és valós adatokon, szürke árnyalatos és színes képeken. A tesztek szerint a (célként nem kitűzött) pontosság is megfelel a korszerű módszerekkel szemben támasztott elvárásoknak. A módszer és az eredmények publikációja: [S1,S2,S3,S5,S6].



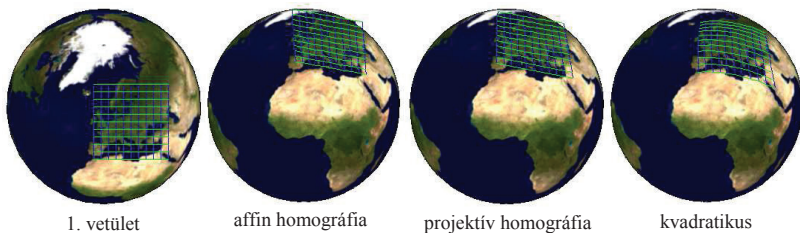
Az Internal Limiting Membrane (ILM) szegmentációja



A Retinal Pigment Epithelium (RPE) szegmentációja

2. ábra: A kidolgozott szegmentáció folyamatának néhány fázisa OCT technológiával készült rácsáló retina képek szegmentációjára.

Az *aktív kontúr* területén végzett kutatásaim célja egy olyan gyors módszer kidolgozása volt, amely lehetővé teszi a valós élekkel nem rendelkező képtartalom-szegmentációt (Optical Coherence Tomography technológiával készült retina képek rétegeinek szegmentációja, 2. ábra). Új eredmény született a szegmentáló görbe menti lokális régiók használatából, amely kombinációja a lokális és a régió alapú módszereknek. Zárt és nyitott görbékre egyaránt használható, lehetővé teszi a szegmentációt vezérlő adatok statisztikai értelmezését teljes régiók tartalmának feldolgozása nélkül. Megadtam egy átlagintenzitás alapján szeparáló Lagrange függvényt, ez alapján felírtam a probléma Normal Flow és Level Set egyenleteit. Javaslatot tettem továbbá az alapmodell kitérinyű továbbfejlesztésére, amelyek növelik az alpmódszer robusztusságát és a rétegek elválasztásának hatékonyságát. Kifejlesztettem egy szoftverkomponenst, amellyel a módszer tesztjeit végeztük. A módszer és az eredmények publikációja: [S7,S10].



3. ábra: Az első kamerakép adott részletéhez tartozó megfeleltetések a második kamera képén. Balról jobbra: lineáris (affín homográfia), projektív homográfia és a kvadratikus transzformációkkal. A megfigyelt objektum implicit felületként adott.

A *3D rekonstrukció* esetében cél volt egyfajta variációs rekonstrukciós módszer megbízhatóságának növelése, alkalmazhatósági tartományának kiterjesztése. A módszer a 3D-be ágyazott felület evolúciójával működik, ahol az evolúciót a (különböző irányokból készült) felvételek képrészleteinek megfeleltetéséből adódó mérték vezérli. Új eredmény a képrészletek közötti kvadratikus transzformáció egyenletei, amelyek mind a leképezési függvényeket, mind a megfigyelt felületet másodrendben (másodrendű invariáns differenciális mennyiségeivel) közelítik. Az egyenletek a vetítési függvényekről nem tételezik fel a lyukkamera modell szerinti alakot. Elvégeztem az eredmények analízisét: az egyes tagok jelentését és viszonyát a projektív és affín homográfiákhoz (3.ábra). Az egyenletek alkalmazásával elvégeztük a variációs rekonstrukciós tesztet, amely igazolta, hogy nagy görbületek esetén a kvadratikus transzformáció megbízhatóbb eredményeket szolgáltat. Fontos megjegyezni, hogy a kvadratikus transzformáció egyenletei általánosabban is használhatók, a gépi látás minden olyan területén, ahol a képrészletek megfeleltetése kulcskérdés. A módszer és az eredmények publikációja: [S4,S8], benyújtva: [S9*].

IV. Tézisek

Az értekezésben a variációs számítás alkalmazásának példáit láthattuk a gépi látás néhány fontos területén. Ezekhez kapcsolódnak az értekezés tézisei.

Tézis 1: A variációs keresztkorrelációs optikai áramlás egyenletei és alkalmazásuk

1.1 Bevezettem a normalizált keresztkorrelációs adatagot szürke árnyaltos és színes képekre variációs keretek között. Levezettem a lokális integrál Euler-Lagrange egyenleteit, a lokális integrál egyenletei alapján felírtam a normalizált keresztkorrelációs funkcionál Euler-Lagrange egyenleteit.

1.2 Kidolgoztam a normalizált keresztkorrelációs adatagot tartalmazó optikai áramlási egyenletek gyakorlati alkalmazásához a közelítő, linearizált numerikus egyenleteket, ehhez első lépésként az analitikus egyenletek kisméretű korrelációs ablakra vonatkozó közelítő formuláját határoztam meg. A közelítő analitikus egyenletből kiindulva kidolgoztam a linearizált numerikus egyenleteket.

1.3 A tézisben ismertetett eredmények validálására és gyakorlati alkalmazására szoftverkomponenst fejlesztettem, amellyel elvégeztem az intenzitásvátozás-türési és numerikus pontossági teszteket a szakirodalomból ismert követelmények szerint.

Tézis 2: Lokális régió alapú aktív kontúr bevezetése, javaslat Lagrange függvényre, a használhatósági tartomány kiterjesztése

2.1 Bevezettem a görbe menti lokális régiók fogalmát szegmentációs célra, ezáltal lehetővé vált a képjellemzők statisztikai értelmezése nyílt és zárt görbékre egyaránt. Javaslatot tettem a lokális régiók szétválasztását lehetővé tevő statisztikai értelmű Lagrange függvényre.

2.2 Két irányban továbbfejlesztettem az alapmodellt. Először, másodrendű görbeillesztéssel lehetővé vált nagy görbületű részek pontos közelítése, ezáltal a lokális integrálási régió méretének növelése a szeparáló görbe mentén. Másodszor, definiáltam az optimális méretű (alakú) integrálási tartományt, amely maximalizálja a lokális régiók elkülönítésének mértékét, növelve a módszer precizitását. Bemutattam, hogy ez utóbbi

probléma lokális variációs számítási probléma. Javaslatot tettem a továbbfejlesztett modell statisztikai értelmű Lagrange függvényére.

2.3 Felírtam a modellekhez tartozó Euler-Lagrange egyenleteket és a Level Set egyenleteket. Kifejlesztettem az eredmények gyakorlati alkalmazását lehetővé tevő szoftverkomponenst. A szoftver szolgáltatja az eredményeket gyakorlati példára alkalmazva, az előszegmentálási módszer eredményeinek javulása állapítható meg, szakértői szegmentálási eredményekre támaszkodó összehasonlításban.

Tézis 3: Képrészletek közötti kvadratikus megfeleltetés (transzformáció) formulájának levezetése, az eredmény megadása invariáns mennyiségekkel

3.1 Felírtam a képrészletek közötti lineáris transzformációt invariáns mennyiségekkel, ezek a vetítési függvények gradiensét és a megfigyelt felület normál-egységvektorát tartalmazzák. Levezettem a képrészletek közötti kvadratikus transzformáció egyenleteit paraméteres formában. Levezettem a kvadratikus transzformáció egyenleteit invariáns formában.

3.2 Megadtam a kvadratikus transzformáció mennyiségeinek kiszámítását lehetővé tevő gyakorlati számítási lehetőségét konstrukcióval, amely pl. véges elem módszerekhez használható. Megadtam a kvadratikus transzformáció mennyiségeinek kiszámítását fix térbeli rácson a Level Set módszerekhez. Az eredményeket alkalmazhatóságát egy többkamerás 3D rekonstrukciós módszer implementálásával, a módszerrel végzett összehasonlító teszttel ellenőriztük.

3.3 A kvadratikus transzformáció analízisével tisztáztam az affin és projektív homográfiákkal való kapcsolatát, továbbá egy olyan alkalmazáson keresztül igazoltuk hasznosságát, ahol a szokásos affin és projektív homográfiáknak kedvezőtlen (nagy görbületű részletek, gyéren textúrázott modell) input adatok álltak rendelkezésre. A kvadratikus transzformáció lehetővé teszi a képrészletek tartományának kiterjesztésével a megfeleltetések pontosságának, és így az erre alapozó módszerek robusztusságának növelését.

A szerző publikációi

- [S1] Molnár József, Csetverikov Dmitrij: "Kereszt-korrelációs optikai áramlás variációs sémája: megvilágítás-változásra invariáns egyenletek", Proc. KÉPAF 2009: 7th Conference of Hungarian Association for Image Processing and Pattern Recognition, CD, Budapest, 2009.
- [S2] J. Molnár and D. Chetverikov: "Illumination-robust variational optical flow based on cross-Correlation", Proc. 33rd Workshop of the Austrian Association For Pattern Recognition, Stainz, Austria, 2009, pp.119-128.
- [S3] S. Fazekas, D. Chetverikov, and J. Molnár: "An implicit non-linear numerical scheme for illumination-robust variational optical flow", Proc. British Machine Vision Conference 2009.
- [S4] J. Molnár, D. Csetverikov: "Másodfokú közelítés implicit felületek síkbeli leképezésére", Proc. Fifth Hungarian Conference on Computer Graphics and Geometry, Budapest, pp. 118-124, 2010.
- [S5] D. Chetverikov, J. Molnár: "An experimental study of image components and data metrics for illumination-robust variational optical flow", Proc. International Conference on Pattern Recognition, Istanbul, pp. 1694-1697, 2010.
- [S6] J. Molnár, D. Chetverikov, and S. Fazekas: "Illumination-robust variational optical flow using cross-correlation", Computer Vision and Image Understanding, vol.114, pp.1104-1114, 2010.
- [S7] J. Molnár, D. Chetverikov, D. Cabrera DeBuc, Wei Gao, and G.M. Somfai: "Segmentation of rodent retinal OCT images", Proc. KÉPAF 2011: 8th Conference of Hungarian Association for Image Processing and Pattern Recognition, Szeged, 2011, pp.140-154.
- [S8] J. Molnár and D. Chetverikov: "Multiview Reconstruction Using Refined Planar Mapping of Implicit Surfaces", Proc. KÉPAF 2011: 8th Conference of Hungarian Association for Image Processing and Pattern Recognition, Szeged, 2011, pp.221-232.
- [S10] J. Molnár, D. Chetverikov, D. Cabrera DeBuc, Wei Gao, and G.M. Somfai: "Layer extraction in rodent retinal images acquired by Optical Coherence Tomography", Machine Vision and Applications. Accepted for publication. DOI: 10.1007/s00138-011-0343-y. 2011.

Bírálat alatt:

- [S9*] J. Molnár, D. Chetverikov: "Quadratic Transformation for Planar Mapping of Implicit Surfaces", Journal of Mathematical Imaging and Vision